Half tidy

Welsh use of the English word '*tidy*' means that something is really rather good. Clearly the term for something that is excellent must be '*half-tidy*'. The state of half-tidyness is quite difficult to attain. Those ferreting around for evidence encounter the half-tidy only rarely, and the tidy infrequently. All too often we have to make do with the downright untidy and try and make something of it.

Systematic reviews can be good or bad, and the studies they review can be good or bad. If it were like Mendel and his peas, then we might have one chance in four of a good review of good trials. This month **Bandolier** has one example of good reviews of good trials, and several heterozygotes. The systematic review process lets us down, usually because the building blocks for the reviews are inadequate.

But we counterpoise this disappointment with some half-tidy observational studies, predominantly from the United States. For instance, a Californian study of glucose self monitoring in 25,000 diabetics is counterpoised with a systematic review of randomised trials that assembled information only 500, and those from inadequate trials. Half-tidy stuff too on troponin tests for cardiac damage, again from the United States.

The other side of the hill

Bandolier sometimes despairs of healthcare industries. Evidence should be their meat and drink, yet they shrink from helping us by giving us evidence the way we want it, in unbiased, independent, systematic reviews. Though there exceptions, too often they are afraid of what's on the other side of the hill. So after incredibly expensive research, carried out over a whole afternoon, **Bandolier** has come up with the answer. Anyone from industry bearing an unbiased, independent, systematic review of a company's product will be rewarded with a **Bandolier** coffee mug. That should do it!

WAITING, QUALITY AND OUTCOME

Waiting times for surgical procedures is not just a British preoccupation. While there are places in the world where operations deemed necessary can seemingly be performed immediately, in many other places patients have to wait for their operation. One such is Canada, where a study demonstrated worse outcomes for patients who have to wait for coronary artery bypass grafts [1].

Study

The study looked at all patients in three Montreal hospitals newly diagnosed as needing CABG over one year ending in December 1994. After some obvious exclusions (like patients needing immediate urgent surgery), there were 266 such patients.

They were interviewed at the time of their enrolment on the waiting list, immediately before surgery, and six weeks and six months after surgery. As well as demographic information and information about their disease severity, symptoms and other medical conditions, the SF-36 was used for quality of life. In additions, major medical events occurring before and after surgery (like myocardial infarction, stroke, death) were recorded from chart review.

Analysis was according to whether patients had their operation before or after the median number of days on the waiting list. Those with a short waiting time waited less than 97 days, the median waiting time. Those with a long waiting time waited more than 97 days.

Results

The average age was 62 years, and 80% were men. The two groups were virtually identical demographically and in disease severity. About 40% had had a prior myocardial infarction.

Five million people a year can't be wrong: **www.ebandolier.com**

Table 1: Major adverse events after CABG according to length of preoperative wait

Major postoperative event	Percent		Relative risk (95% CI)	NNH (95% CI)
	Fewer than 97 days	More than 97 days		
Stroke	5	11	2.4 (0.9 to 5.9)	
Myocardial infarct	2	5	3.5 (0.8 to 17)	
All bad events	**11**	**24**	**2.3 (1.3 to 4.1)**	**7.3 (4.4 to 21)**

There were no differences in SF-36 results at baseline. Patients with longer waits had significantly worse scores for several of the eight domains of SF-36 immediately before and six months after the operation, including *physical functioning, vitality* and *general health* at both times. Six months after operation *physical role* and *mental health* were also significantly worse in patients who had waited more than 97 days than in those who had waited less. Improvements in these domains that occurred after operation in patients with shorter waiting times just did not happen in those who waited longer.

There was no difference between the groups in the occurrence of major medical events before surgery. After surgery, patients with the longer wait had significantly more major medical events (24%) than those with a shorter wait (11%) (Table 1). For every seven patients who waited longer than 97 days, one more had a major medical event postoperatively than if they had a wait of less than 97 days.

By six months after surgery where the waiting time was less than 97 days, 17 of 20 employed patients (85%) remained employed. Where the waiting time was more than 97 days, 10 of 19 employed patients (53%) remained employed.

Comment

Studies on the effect of waiting time for surgery appear to be rare. It's hardly something that an ethics committee would approve for a randomised trial, so observational studies will be the only form of evidence we get. The strengths of this one were that it was comprehensive (all patients in Montreal for one year), thorough, and by chance the two groups were homogeneous at baseline. What was odd was that it took six years to get to print.

The results will surprise no-one, but rather confirm our observations that ill people go downhill if effective treatment is not forthcoming. It perhaps serves to concentrate the mind on waiting time (all of it, including waiting for a consultant appointment), and remind us that effective management is as important as effective interventions.

References:
1   J Sampalis et al. Impact of waiting time on the quality of life of patients awaiting coronary artery bypass grafting. Canadian Medical Association Journal 2001 165: 429-433.

MORE EVIDENCE ON CHRONIC HEADACHE

Chronic headache is a big problem for individuals, their professional carers and health services. Reviews examining the efficacy of treatments are welcome. Two reviews, of very different type, inform on chronic headache treated with antidepressants [1], a common condition, and cluster headache [2], much more rare.

Antidepressants for chronic headache

This thorough review from the USA [1] did extensive searching for randomised trials comparing antidepressant prophylaxis with placebo for chronic headache. A range of outcomes were extracted, including global improvement (usually at least 50% reduction in headache burden, itself obtained by multiplying frequency of headaches by their severity), headache burden, and use of analgesics.

Results

There were 38 randomised trials with useful data, predominantly in chronic migraine and chronic tension headache. Average trial size was 50 patients and the average duration was 10 weeks. More recent studies used International Headache Society classifications for headache diagnosis. Tricyclic drugs were represented primarily by amitriptyline (12/19), serotonin blockers by pizotifen (12/18) and SSRIs by femoxetine (3/7). Dose varied (for instance for amitriptyline between about 10 mg and 150 mg daily; and readers should beware that the notation "qd" in American means once-a-day).

Pooled data for global improvement, headache burden and analgesic use all favoured antidepressant over placebo. For global improvement 31% more patients improved with placebo, with a number needed to treat of 3.2 (95% confidence interval 2.5 to 4.3). With a rate ratio of 2.0, this implies that about 60% of patients using antidepressants have a global improvement approximating to about 50% reduction in headache burden.

There was no observed difference between different classes of drugs. Type of headache, type of antidepressant or duration of treatment did not affect results, nor did several aspects of trial design (parallel group versus crossover), or trial size or trial quality. Adverse events were not described in the review.

Cluster headache

Cluster headache (unilateral, excruciating severe attacks of pain principally in the ocular, frontal and temporal areas recurring in separate bouts with daily or almost daily attacks for weeks to months) is rare. It affects fewer than one person in a thousand in their lifetime. A review [2] looks at all the evidence on cluster headache, including classification, epidemiology, aetiology, pathophysiology and drug and surgical treatment, while another, though not overty systemtic, is comprehensive [3].

It's a fun review, impossible to summarise for **Bandolier**. The best evidence for drug treatment is for subcutaneous sumatriptan.

Comment

Two interesting if different reviews on chronic headache. One question for which the answer remains blurred is when are headaches sufficiently chronic to need prophylactic treatment? Is it one a day, one a week, or one a month? There are two ways of looking at this.

A small proportion of triptan users consume large numbers of tablets, sometimes in excess of one a day. They represent a significant minority of triptan users and triptan use, and clearly need expert help.

Another view is that even one headache a month, that causes significant disability and for which triptans are not effective, should be a trigger for prophylactic therapy.

Whatever, prophylactic treatment of chronic headache is a growth area where significant change can be expected. Watch this space.

References:
1   GE Tomkins et al. Treatment of chronic headache with antidepressants: a meta-analysis. American Journal of Medicine 2001 111: 54-63.
2   JM Zakrzewska. Cluster headache: a review of the literature. British Journal of Oral and Maxillofacial Surgery 2001 39: 103-113.
3   Dodick DW et al. Cluster headache. Cephalalgia 2000 20: 787-803.

GLUCOSE SELF-MONITORING AND GLYCAEMIC CONTROL

Diabetics with type 1 diabetes using insulin are recommended to test their blood glucose levels several times a day. In the USA type 2 diabetics using insulin or oral hypoglycaemic agents are recommended to test their blood glucose at least daily. There are no recommendations about monitoring for type 2 diabetics using diet and exercise to control their glucose levels.

Glucose self-monitoring is not without cost, and the annual cost of strips for someone monitoring their glucose level three times a day is about $850 in the USA and £330 in the UK. The efficacy of glucose self-monitoring controversial. There are concerns over its effectiveness even for type 1 diabetics, let alone type 2 diabetics. A new study from the Kaiser Permanente diabetes registry in Northern California covering about three million people [1] may help.

Study

The registry collects information about diabetics, including prescribed drugs and control strips, outpatient and hospital visits, and is estimated to have information on 96% of all diabetics in the programme. The study set out to examine the average daily strip use by diabetics throughout 1996, and relate that to the first haemoglobin A1c measurement in 1997 as a measure of glycaemic control. At the time of the study there were no Kaiser Permanente practice guidelines for self-monitoring.

Information from patients about themselves and their diabetes was collected by questionnaire or telephone interview, including information about treatment and use of diet and exercise. Automated pharmacy records were used to calculate glucose strip use, and linked records to evaluate emergency room and hospital admissions.

Table 1: Definition of average daily reagent strip use from prescription analysis

	Strips a day
Three times daily or more	≥2.5
At least daily	<2.5 but ≥0.75
Occasional	<0.75 but >0
No self monitoring	0

The American Diabetes Association 1997 guidelines were for self monitoring three times a day for type 1 diabetics, and daily for type 2 diabetics treated with insulin or oral hypoglycaemic agents. These rules were used to define adherence and non-adherence. Definitions of average daily reagent strip use from prescriptions are shown in Table 1.

Haemoglobin A1c results were corrected for demographic differences between different categories of strip use. Information on excluded patients was also given and results were similar to those in the main study.

Results

There were just under 49,000 diabetics with full membership of the health programme for the whole of 1996 and 1997, and 23,412 who additionally responded to the survey, in whom type of diabetes could be determined and in whom haemoglobin A1c was measured in 1997.

Type 1 diabetics

There were 1160 such patients, of whom 395 measured blood glucose on average three times a day. Their haemoglobin A1c was, on average 1.2% lower (7.6%) than those who

Figure 1: Average 1997 haemoglobin A1c values according to diabetes diagnosis and average daily glucose strip use during 1996
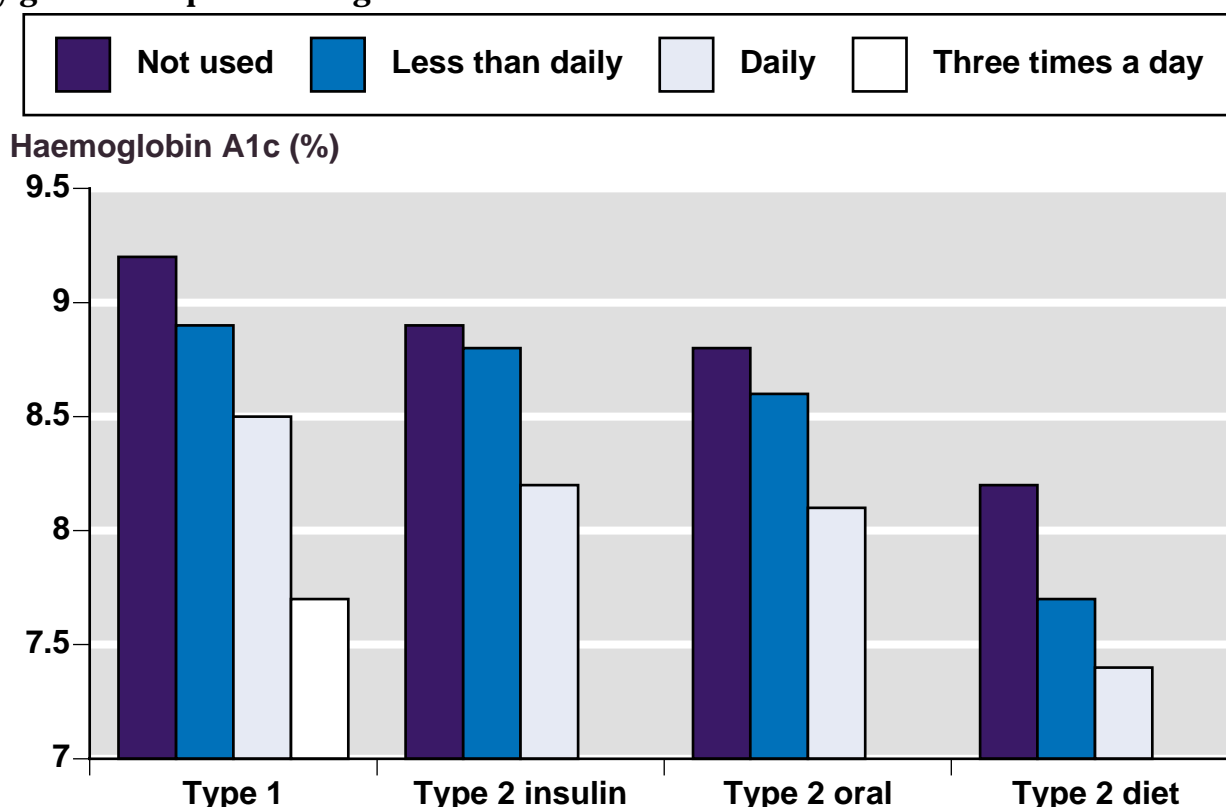
monitored less frequently (8.8%). Adherent diabetics who monitored three times a day tended to be a few years older and more likely to be women than non-adherent diabetics. More monitoring was associated with lower levels of haemoglobin A1c (Figure 1).

Type 2 diabetics

Seven thousand five hundred of 23,000 patients measured their glucose at least daily. Haemoglobin A1c in diabetics treated with insulin or oral hypoglycaemic agents was 0.7% lower (8.1%) than in those who monitored less frequently (8.8%). Adherent diabetics who monitored once a day tended to be older and were more likely to be women than non-adherent diabetics. More monitoring was associated with lower levels of haemoglobin A1c (Figure 1).

Adherent type 2 diabetics were more likely than non-adherent diabetics to attend for eye examination and use diet and exercise as part of their therapy. They also had more frequent emergency room attendance (31% vs 26%) and hospital admissions (17% vs 13%), but we don't know if these were for hypoglycaemic episodes.

Comment

Evidence for the benefits of glucose self-monitoring, especially in type 2 diabetics, is lacking. The paper [1] has a good reference list for the evidence that exists. Some randomised studies have been done, and there is even a good meta-analysis [2], but a good meta-analysis of poor trials. The total number of patients for analysis in randomised trials was 558, dwarfed by this study, and with high dropout rates and other methodological problems. New randomised trials are unlikely, despite the apparent equipoise.

So observational studies assume increasing importance, especially when, like this, they are large and detailed. The result, that increased self monitoring up to recommended levels is associated with better glycaemic control, is the expected one. Unfortunately such trials do not allow us to conclude that increased self-monitoring results in better glycaemic control, though that might seem obvious.

For type 2 diabetics in particular there were interesting differences between those patients who measured their blood glucose daily and those who did not. This limits the applicability of the study, and the authors themselves conclude that this is only supportive evidence in favour of the benefits of self-management in type 2 diabetics.

This was a relatively short-term study. *Bandolier* 84 described a long-term study from Washington State where diabetic patients who achieved a 1% fall in haemoglobin A1c reduced their overall healthcare costs by almost $1,000 a year, and sustained it, but only after the second year.

These large, real-world database studies emerging from United States healthcare providers often have important academic affiliations. This study [1] had NIH and American Diabetes Association support. The motives of better health care and lower cost are often mutually supportive.

References:
1   AJ Karter et al. Self-monitoring of blood glucose levels and glycaemic control: the Northern California Kaiser Permanente diabetes registry. American Journal of Medicine 2001 111: 1-9.
2   S Coster et al. Self-monitoring in type 2 diabetes mellitus. A meta-analysis. Diabetic Medicine 2000 17: 755-761.

TROPONIN TESTS FOR CARDIAC DAMAGE

One of the problems with new diagnostic tests is getting one's brain around what they mean and how and when to use them. There will be masses of papers that are completely useless, where test results are compared in really sick patients with those with no disease. We know that such studies are hopelessly biased (***Bandolier*** 70). What we need are more real world examples of how and when tests can be used to advantage.

For some years new tests of damage to cardiac muscle have been available, measuring proteins called troponins. There are two varieties, T and I, and different forms of tests are used with different levels of sensitivity. There's little doubt that they *look* useful. Hard evidence has been difficult to come by. Three new publications help a lot.

Troponins in acute coronary syndromes

The evaluation of patients with chest pain suggestive of heart attack is a common and costly problem. We know that patients with chest pain and elevated ST-segments of an electrocardiogram are at risk and should be treated. A meta-analysis from California [1] tells us that patients with non-ST elevations but with a positive troponin result are also at increased risk.

Review

The review sought clinical trials and cohort studies evaluating patients with suspected myocardial infarction. Excluded were studies only on patients with infarction, case-control studies, articles not reporting mortality and those that included patients with ST-elevations. Outcomes chosen pre hoc were mortality, and the combined end point of death or infarction for comparison between patients who had at least one positive troponin test and those patients whose troponin test was negative. Where more than one time point was used, information closest to 30 days was chosen.

Results

After exclusions, seven clinical trials and 19 cohort studies were used, with 5,400 patients examined with troponin T and 6,600 with troponin I. The mean age of patients was 63 years, and two thirds were men. About 25-50% of patients had previous myocardial infarction. Patient cohorts included acute coronary syndrome and unstable angina.

Results for troponin T for death are shown in Figure 1 and for death or myocardial infarction in Figure 2. Outcomes for patients with a positive test occurred more frequently than in patients with a negative test. Similar results were obtained with troponin I.

Rates of death were about 1.5% with a negative troponin test and 6% with positive tests. Rates of death or myocar-

Figure 1: Death after positive or negative troponin test
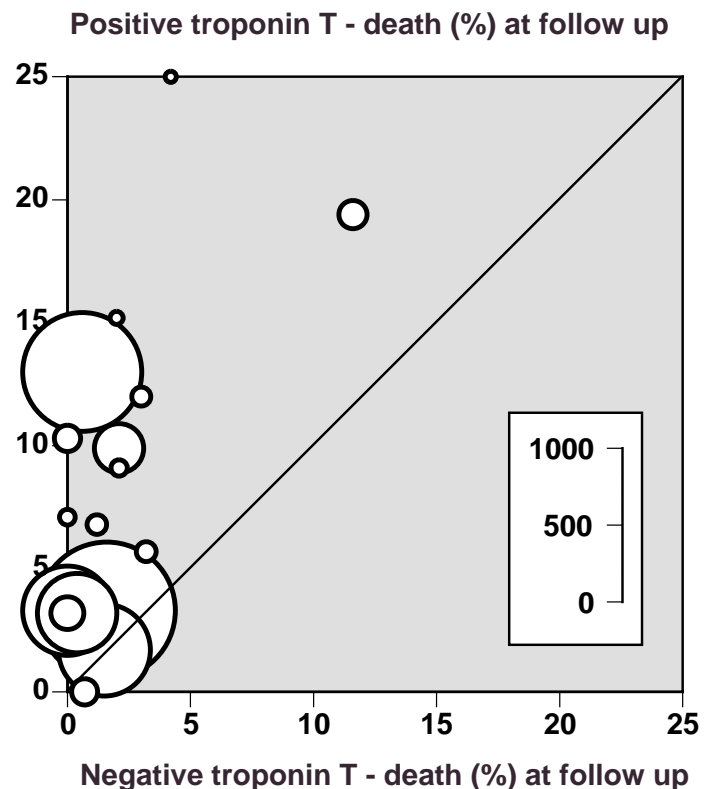
Figure 2: Death or infarction after positive or negative troponin test
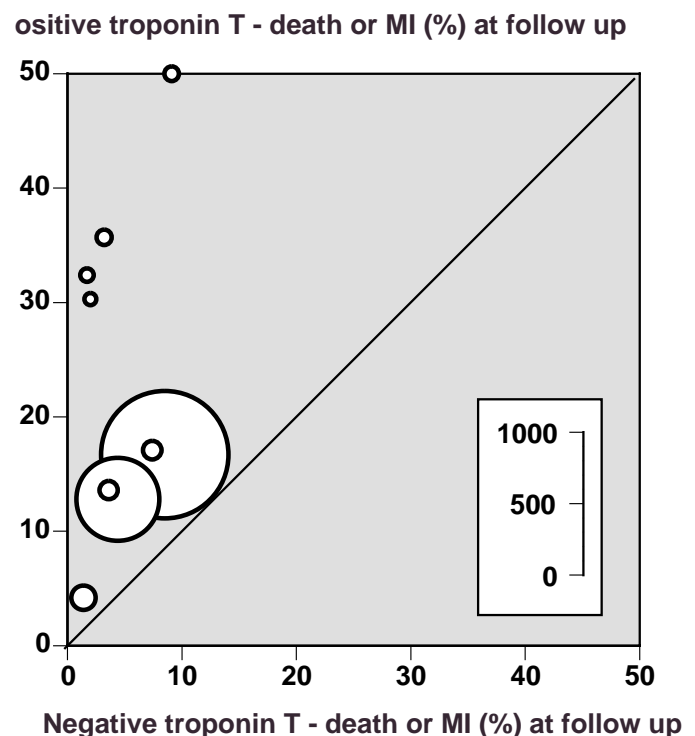
dial infarction were about 5% with a negative troponin test, and about 16% with a positive troponin test. The relative risk was between 3 and 4 (Table 1).

Troponins and cost

Barriers to introducing new tests are not restricted to lack of evidence of efficacy. New tests, like new drugs, are more

Table 1: Outcomes in patients suspected of myocardial infarction without ST-elevation on electrocardiogram

Test and outcome	Number/Total (%)		Relative risk (95%CI)
	Positive test	Negative test	
Troponin T			
Death	99/1635 (6.1)	53/3524 (1.5)	4.1 (2.9 to 5.7)
Death or MI	143/872 (16.4)	85/1412 (6.0)	3.0 (2.3 to 3.9)
Troponin I			
Death	108/1981 (5.5)	77/4422 (1.7)	3.3 (2.5 to 4.4)
Death or MI	13/51 (25.5)	12/240 (5.0)	4.9 (2.4 to 10)

expensive than the ones they replace, and budgetary constraints frequently hamper introducing new tests even when there is abundant evidence. Good evidence that new tests can reduce overall costs is usually lacking, but for troponins a randomised study from Connecticut concludes that their use can reduce overall costs [2].

Study

Patients presenting to the emergency department consenting to the study were randomised (method not stated) to receive a standard clinical evaluation with electrocardiogram and CK-MB determinations, or the same package but with serial troponin T evaluations at presentation, and three and 12 hours after presentation. Exclusions were ST-elevation at presentation, defibrillation or resuscitation.

Clinical data were collected, with information about subsequent care, length of hospital stay, and total charges. Patients were followed up to 30 days. A pre-specified subgroup analysis was by presence or absence of acute coronary syndrome (confirmed unstable angina or acute myocardial infarction), and the final diagnosis was made by a cardiologist using WHO criteria and unaware of troponin values.

Results

The groups were well matched at baseline. The mean age was 65 years and half were women. Serial troponin T measurements were used in 409 patients, but not in 447.

Length of stay was significantly shorter for patients who had troponin measurements with acute or non-acute coronary syndrome (Table 2). Total hospital charges were lower for patients who had troponin measurements, significantly so for those with acute coronary syndrome (Table 2). On average, total costs for patients with troponin measurements were about $900 less, representing a potential annual saving to the hospital in Bridgeport of about $4 million.

For a diagnosis of myocardial infarction a positive test in serial troponin measurements had a positive likelihood ratio of 17 and a negative likelihood ratio of 0.04. The overall rate of myocardial infarction was 6%, so that the post-test probability of myocardial infarction with a positive troponin test alone was about 45%, while the post-test probability with a negative test was about 0.2%.

Algorithms and outcomes

So we know that elevated troponins are associated with bad things happening in people without ST-elevations, and we know that using troponins to evaluate people attending emergency departments with suspected heart attacks can save money. The cynics among us will then say that we need evidence about how the new tests can best be incorporated into work patterns in busy emergency departments. For instance, can the laboratory produce results within the time needed? If the results are not available to clinicians making decisions about whether to admit a patient, then why bother? A study from San Diego [3] claims that they can form part of a 90-minute triage using point of care testing.

Table 2: Use of troponin tests on resource use in hospital for patients with non-acute and acute coronary syndromes

Diagnosis	Outcome	With troponin measurement	Without troponin measurement
Nonacute coronary syndrome	**Stay (days)**	**1.2**	**1.6**
(N=654)	Charges (US$)	4,487	6,187
Acute coronary syndrome	**Stay (days)**	**3.7**	**4.6**
(N=202)	**Charges (US$)**	**15,004**	**19,202**
Shaded areas indicate statistically significant reduction with toponin			

Figure 3: Destination of patients after 90-minute triage

Patient destination (% of total)

Legend: Baseline, Algorithm
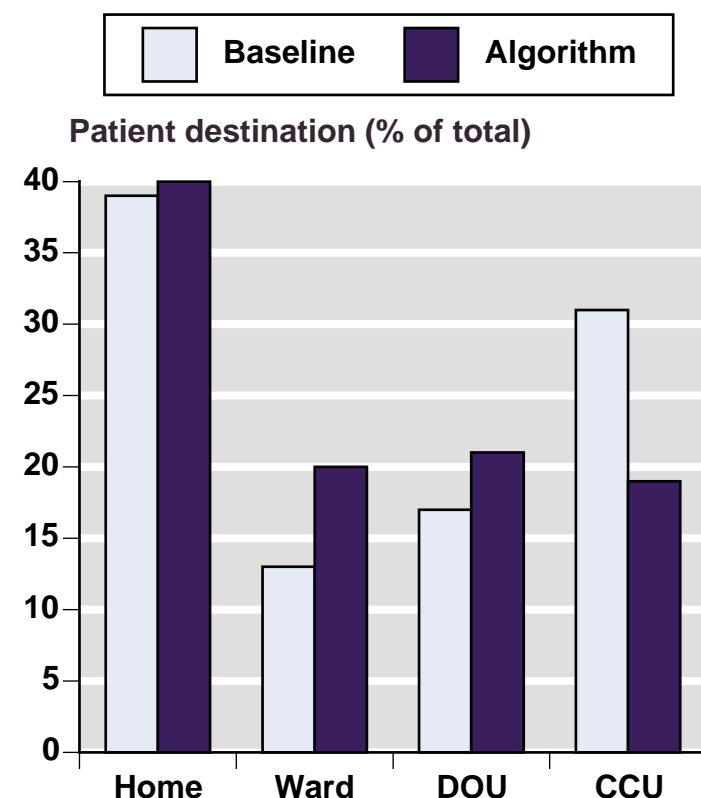
Study

This study tested a cardiac marker algorithm in 1,285 consecutive patients presenting to an emergency department with symptoms of cardiac ischaemia over nine months to April 1999. The pathway cannot be described in detail here, but involved electrocardiograms, clinical examination and judgement, and the use of troponin I, CK-MB and myoglobin tests done using a point of care testing machine at presentation and 30, 60 and 90 minutes later. All MI diagnoses were confirmed by a cardiologist using WHO criteria. The destination of patients (CCU, direct observation unit or DOU, ward or home) was compared with a six month period to March 1997.

Results

The mean age was 64 years and 98% were men. Of the 1,285 presenting patients, 66 (5%) had a final diagnosis of myocardial infarction. All of these were diagnosed within 90 minutes by use of the cardiac marker algorithm and/or electrocardiograph changes.

Of the 508 patients discharged to home, 90% were discharged within 90 minutes and all within six hours. Thirty-day follow up of these patients showed that one (0.2%) was readmitted with myocardial infarction, and 7 (2%) were readmitted with unstable angina.

Of the 771 patients admitted to hospital, 32% were admitted to CCU, 35% to a direct observation unit, and 34% to a ward. Comparison of patient destination with a previous period showed a 40% reduction in the use of CCU beds (Figure 3).

Comment

Bandolier's ideal is to have the triple whammy of evidence of effectiveness, evidence about costs, and evidence about how to make a change. These three papers together come a long way to fulfilling those requirements for the use of troponin tests. They demonstrate that patients without ST-elevations but with positive troponin results have increased risk of death or myocardial infarction, that use of troponins has the capacity to reduce hospital resource use, and that people have figured out how to aggressively triage patients with suspected heart attack and come up with the right answer. Reducing CCU bed use by 40%, if replicated elsewhere, would appear a pretty stupendous result for hard-pressed hospitals with staffing problems.

Bandolier is conscious that in making a précis of these three papers much detail has been omitted. What is terrific is to have this quality and quantity of information about a laboratory test. It has been too easy in the past to criticise laboratory tests for inadequacy of evidence. This begins to put it right.

What is particularly interesting is that the three studies use different architectures to investigate their questions. One is a classic meta-analysis of clinical trial and cohort studies, but using the characteristic of good diagnostic studies that patients have to be consecutive to avoid bias. Another is a randomised trial to assess health economic outcomes. And the third tests out a diagnostic algorithm, again on consecutive patients.

Bandolier is rather impressed, especially because the diagnostics "industry" has lagged far behind therapeutics in the quality of the evidence it believes necessary to justify its existence.

References:
1   PA Heidenreich et al. The prognostic value of troponin in patients with non-ST elevation acute coronary syndromes: a meta-analysis. Journal of the American College of Cardiology 2001 38: 478-485.
2   S Zarich et al. Impact of troponin T determinations on hospital resource utilization and costs in the evaluation of patients with suspected myocardial ischaemia. American Journal of Cardiology 2001 88: 732-736.
3   S Ming et al. Ninety-minute accelerated critical pathway for chest pain evaluation. American Journal of Cardiology 2001 88: 611-617.

HOW SYSTEMATIC REVIEWS CAN DISAPPOINT

Bandolier tries to find examples of systematic reviews where there is a solid take-home message. After all, we are familiar with uncertainty, and systematic reviews that merely bleat, often inadequately, about inadequacies of research are rather depressing. The trouble is that people use phrases like "evidence-based medicine", or "meta-analysis" or "systematic review" as some form of talisman. Attach one of these phrases to a point of view and an argument is won!

That fails to take into account the fact that reviews can be awful, and even completely wrong. Often systematic reviews try to cover too much ground, and end up giving us too superficial a view of a problem. And even the best of reviews of good studies can leave us in the lurch when there are not enough good studies.

In consequence, here are some examples of systematic reviews that don't help for one reason or another, and on important topics where an answer would be most helpful.

Urge urinary incontinence

A big problem this, and one where we'd like some answers, please. Finding a meta-analysis [1] might make the morning brighter, but reading it made the afternoon stormy.

The review did all the right searching for studies comparing tolteridone versus oxybutinin in the treatment of urge urinary incontinence. There's a clue to a problem, because the review didn't set out to tell us that the treatments worked or how well they worked, but only if one was better than another. Fuzzy thinking from the start.

There were four trials (but two were abstracts). How many patients in these four trials? We have no idea, because we are not told. There may have been some differences favouring oxybutinin for incontinent episodes or tolterodine for voided volume per micturition, and some adverse event data in favour of tolterodine, but without baseline data and numbers it is completely uninterpretable.

This might be the worst meta-analysis ever from a reader point of view, but it badges itself as having been done using "Cochrane methodology", the ultimate designer label for meta-analysis.

Chronic fatigue syndrome

A review of interventions for chronic fatigue syndrome [2] from York and San Antonio is excellent in many ways. It is beautifully done, with great methods, includes 44 controlled trials (some randomised) and tells us a fair bit about the trials. The problem here is that the 44 studies covered an enormous range of interventions (exercise, drugs, supplements, complementary) plus a wide range of outcomes (psychological, physical, quality of life, physiological and laboratory). If there's a statistical difference, the study gets a tick.

To be fair to the authors of the review, what they have done is what most folk would do. But is it enough? For instance, most studies were small (average about 60 patients, median many fewer). If statistical significance was set at 1 in 20 (5%, 0.05), then that is no great shakes. What would it have looked like at 1%, or 0.01? And what *is* a useful outcome for chronic fatigue, anyway?

So where's the beef? Let's face it, a significant result in a randomised trial of massage therapy in five (yes, 5, or V) patients doesn't make us sit up and beg for more, especially when we've no idea what benefit means. For what it's worth, behavioural intervention trials were of higher validity and were more often positive.

Do decision aids help decision making?

This review [3] focused on studies in which real patients made actual decisions. The review methods were terrific, the description of observational studies and randomised controlled trials excellent and the studies themselves varied but were often large and asked the important question. Do decision aids influence the decisions that patients take?

There were about 20 studies. The bottom line was that patient decision aids do not influence the decisions that patients take.

But there's more than just this stark answer, and for people involved with patient decisions or interested in using decision aids for patients, this is a wonderful resource. It enlightens about research that has been done, and thinks about research that might be done.

Comment

Three systematic reviews and meta-analyses, all on important topics. A mix of quality. What is important is not that they *are* systematic reviews, but how *good* they are, and how well they inform. Systematic reviews need brains as well as stamina.

People doing systematic reviews to get another paper on their curriculum vitae or some academic brownie points can be content with useless reviews. If the reviews are about clinical impact rather than about impact factors, then they need to be useful. Can they be used in Pontycymer on a wet Thursday afternoon? That's the ultimate test of quality.

References:
1 MA Harvey et al. Tolteridone versus oxybutinin in the treatment of urge urinary incontinence: a meta-analysis. American Journal of Obstetrics and Gynecology 2001 185: 56-61.
2 P Whiting et al. Interventions for the treatment and management of chronic fatigue syndrome. JAMA 2001 286: 1360-1368.
3 C Eastbrooks et al. Decision aids: are they worth it? A systematic review. Journal of Health Services Research & Policy 2001 6: 170-182.